Trusted Tiny Things: Making Devices in Smart Cities More Transparent

Stanislav Beran, Edoardo Pignotti, and Peter Edwards

Computing Science & dot.rural Digital Economy Hub , University of Aberdeen, Aberdeen AB24 5UA, UK
{s.beran,e.pignotti,p.edwards}@abdn.ac.uk

Abstract. In this demo we present the *Trusted Tiny Things* system that can be used to interrogate Internet of Things (IoT) devices and present users with information about their characteristics and capabilities. The system consists of a mobile application used to retrieve information about IoT devices supported by RESTful web services. In order to infer IoT device capabilities our services perform reasoning over the provenance of devices characterised using a number of Semantic Web technologies. In this demo we illustrate the use of the system with two distinct IoT devices: an NFC tag used at bus stops to provide a means to access real-time bus timetables, and a black box device installed into vehicles by insurance companies to track driving behaviour.

Keywords: internet of things, provenance, transparency

1 Introduction

The vision of the Internet of Things (IoT) is a dynamic global network based on standard and interoperable communication protocols where physical and virtual 'things' have identities, physical attributes, and capabilities and are seamlessly integrated into the existing internet infrastructure. The IoT is thus built upon a range of sensors and other devices [1] that together represent the 'things'. These devices range from passive radio tags to internet connected sensor platforms and embedded computers. Deployments of such devices are increasingly commonplace in Smart City environments to capture, analyse and exchange streams of information. For example, passive NFC (Near Field Communication) tags are currently in use by Aberdeenshire Council in Scotland UK to provide smartphone access to timetable information for a particular bus stop. Active IoT devices include in-car black boxes [2] being introduced by insurance companies to assess the behaviour of drivers and affect their premiums. Such devices and their associated entities (e.g. web services or other devices in the IoT ecosystem) could potentially generate vast amounts of data. This data may contain personal or any other confidential data that users may not wish to share. Alternatively, user may wish to know how the data is used and by whom.

Questions that a user might ask include: *What kind of data does the thing collect? Is the data transmitted? If so, how and to whom? For what purposes are*

the data used? What control do I have over any aspects related to the generation and use of this data?. These questions are reflected in the "TRUSTe Internet of Things Privacy Index - GB Edition[1]" study where more than 80% of the 2,005 people interviewed were concerned about such issues. Our proposed solution is largely based on participatory design activities we conducted during the course of the project. To date, we have conducted a number of participatory design events involving a total of 77 participants with different technological backgrounds[3].

These questions are particularly relevant in the context of Smart City development. Bartoli et al.[4] identified privacy and sensitive data management as one of the key issues to be addressed during the design of Smart City systems. The authors discuss the different perspectives to be taken when assessing security or insecurity of a particular service and highlight that service should be configurable to best suit specific user needs based on sensitive data the user provide. Bartoli et al. also state: "The number of users, and the volume and quality of collected data, will also increase with the development of Smart Cities. When personal data is collected by smart meters, smart phones, connected plug-in hybrid electric vehicles, and other types of ubiquitous sensors, privacy becomes all the more important. The challenge is, on the one hand, in the area of identity and privacy management, where, for instance, pseudo-nomination must be applied throughout the whole system, in order to separate the data collected about a user (which is required in order to provide high- quality personalised services) from the users real identity (which is required for purposes such as accounting); this includes that the usage of addressing identifiers, such as IP or MAC addresses, for the purpose of identification must be avoided in future systems."

The Trusted Tiny Things (T3) project is investigating how Semantic Web technologies can be used to describe the context surrounding IoT devices (e.g. manufacturer, owner, security method) and to reason about device capabilities. As 'things' become more interconnected this context should also include provenance information: a record of the entities (devices or services) and processes (data transmission, data analysis, decision making) involved in the creation and use of data. A formal representation of provenance has been identified as essential to support users (and machines) to better understand and trust data[5]. For example, in the car black box scenario, provenance could be used in order to understand what kind of data the box is collecting, what agents or services are using this data, and for what purpose.

In the remainder of this paper we discuss our T3 software system and the supporting semantic framework. We conclude by presenting a list of features being demonstrated about the system through the use of two case studies (in-car black box and bus stop).

[1] https://www.truste.com/gb-internet-of-things-index-2014/

2 The Trusted Tiny Things Architecture

We have developed a software infrastructure and mobile app (see Figure 2) that can be used to query, update and register IoT devices and to notify the user of any changes in the capabilities of a registered device.

2.1 User Groups

There are two user groups concerned with usage of our system.

- **Device owners and manufacturers** - this user group is concerned with registering IoT device into the system and updating device characteristics and provenance records through our Restful API according to our guidelines[2].
- **Users of the IoT system (devices and services)** - These are the users of the IoT device and associated services and they interact with our system through NFC-enabled Android mobile application.

2.2 System Architectural Layers

The system is composed of multiple layers (see Figure 1):

- **Storage layer** - it is used for storing and retrieving device metadata, provenance record and information about users. We are using two different Jena TDB[3] repositories. One repository holds publicly available provenance of IoT devices. The other repository holds a record of user preferences and accepted devices. We also utilise MySql Database to store confidential user data for increased security(e.g smartphone IDs used to identify users, associations with the IoT devices and statistics how users are using our system).
- **Ontological layer** - it contains a number of ontologies used to support the metadata in the Storage layer. These ontologies are discussed in Section 3.
- **Service Layer** - it defines methods for querying, updating, and synchronising data from the devices. It is also capable of notifying users of any changes in an IoT device or service (e.g. a new organisation is using information provided by an existing IoT device).
- **RestFul API Service Layer** - it provides uniform access to the system from external applications such us our Trusted Tiny Things mobile app.

2.3 Core Services

The system consists of five core Java EE based services:

[2] http://t3.abdn.ac.uk/guidelines
[3] http://jena.apache.org/documentation/tdb/

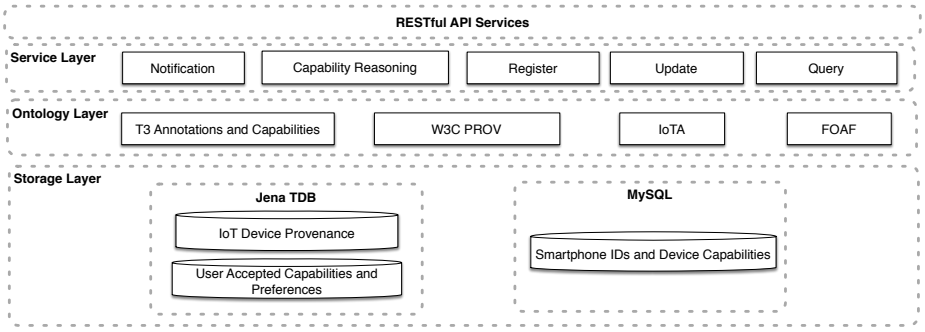

Fig. 1. Trusted Tiny Things System Architecture diagram

- The *Registration* service is used to register new devices into the system. It provides methods for generating links between devices and the physical entities they represent using the IoTa domain model. When user registers device he/she is given unique ID, which should be used over the whole IoT system to send generated or intended provenance data to our system. For example, if the device is communicating with the server and server is acting upon the data that was collected from IoT device, provenance of the transactions or processes that were carried out externally should be captured and send to our system.
- The *Update* service provides methods for updating information about devices in the system. This includes the ability to update the provenance of the device and its use. On how to update provenance and other IoT device information reader can refer to our guidelines. This service also provides a method for associating smartphone IDs and devices when users accept (or decline) the capabilities of the device via our app.
- The *Capability Reasoning and Filtering* service is designed to reason about the provenance information associated with devices in order to infer direct or indirect capabilities. This service makes use of a rule-based reasoner implemented using the TopBraid SPIN API[4] to evaluate rules of the kind described in Section 3.1. Since the smart devices, their services and third-party smart application can generate vast amount of provenance metadata, it was necessary to implement an efficient mechanism to filter redundant repeated capabilities that are already being recorded. For example, public security camera is capable of motion detection. For each motion event activation the camera would generate provenance, which would be sent to our system. A user may only wish to know that camera has motion capabilities.
- The *Query* service is used for extracting information about an IoT device based on the metadata stored in our system (including provenance and capability inferences).

[4] http://topbraid.org/spin/api/

– The *Notification* service checks for changes in the capabilities of registered devices every hour. If the service detects such changes push notification messages are sent to the relevant users (using the smartphone IDs associated with the device) informing them of the new capabilities.

2.4 Smartphone app

The services described above are accessible from our mobile app via a RESTful API layer and JSON[5] is used as the data interchange format. Devices in our API are recognised by a custom URL *http://t3domain/devices/{DeviceID}* where the DeviceID represents the device identifier encoded in our Trusted Thing NFC tag. Using the devices URL we can support different types of GET and POST actions to retrieve or create information about devices. The app listens for NFC events and intercepts the signal in order to determine if the device was registered with our system. If so, it would retrieve characteristics and list of inferred capabilities derived from past behavior of the device. The app also summarizes in pictorial representation all the organizations associated with it and the type of data that are being collected. If user is happy with the capabilities, he/she has the option to either Accept or Decline the use of the smart device or service associated with it. If accepted, user can store the device under his/hers own nickname so the user can reference to it in the future. Once user accepted the device capabilities, he/she also subscribes in our system to be notified about any changes in capabilities (e.g. capability changes or new capability is detected).

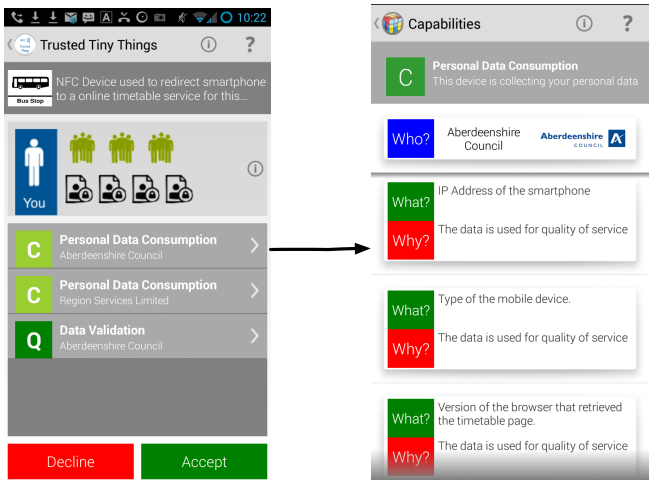

Fig. 2. Landing page (left) and Capability page (right) of the Trusted Tiny Things mobile app.

[5] http://www.json.org/

3 The Trusted Tiny Things Semantic Framework

In order to inform the design of a semantic framework for IoT devices we have conducted three participatory design events involving a total of 14 participants with different technological backgrounds. Participants were asked to discuss issues surrounding the capabilities of IoT devices. Questions were posed such as: *What do you think are the capabilities of this device?* and *What kind of capabilities would you want to be aware of before interacting with this kind of device?*. Guided by user requirements we have designed an ontological framework to support inferences about device capabilities using provenance.

In order to retrieve information about IoT devices (characteristics, provenance, capabilities, etc.) it is necessary to be able to identify things (e.g. in car black box,smart fridge) and their IoT components (tag, device, sensor or service).

Bandara et al. [6] presents a semantic model for describing IoT devices. While this model is capable of describing device characteristics, it does represent relationships between the devices, services and their usage in the context of smart domains (e.g Smart Cities, IoT systems). Kortis et al. [7] describe an ontology that represents knowledge about 'Things' in the IoT domain and the way they should interoperate. The authors have created a model describing IoT concepts by introducing ontological definitions such as Physical Entity, Control Entity, Electronic Device, Smart Entity Cluster and Smart Network. While this ontology is capable of capturing some relationships it is mostly focused on finding a common framework to allow deployment of IoT devices into the existing Internet infrastructure for service discovery and it is not suitable for our needs as it is too focused on low level service description.

The Internet of Things Architecture[6](IoTa) is another project working towards building a common architecture for the future Internet of Things. They have developed a conceptual model [8] to describe the IoT domain based on previous work from Serbanati et al. [9] and Haller [10]. The main aim of the IoTa Domain model is to characterise the different entities in the IoT domain (e.g. User, Service, Device, Physical Entity, Virtual Entity and Resource) see Figure 3.

In order to reason about device capabilities we need to be able to represent the provenance of an IoT device (e.g. processes, agents and the data being used and generated). In our model we describe such provenance using the emerging W3C PROV-O[7] ontology as it is designed to be applicable to a wide range of applications and domains. PROV-O defines concepts such as: *prov:Entity* (physical, digital, conceptual); *prov:Activity* (something that occurs over a period of time and acts upon or with entities.); and *prov:Agent* (something that bears some form of responsibility for an activity). Using this ontology it is possible to describe who is the agent responsible for a specific process (*prov:Activity*) taking place (e.g. who collects my position location data) or what information

[6] http://www.iot-a.eu
[7] http://www.w3.org/TR/prov-o/

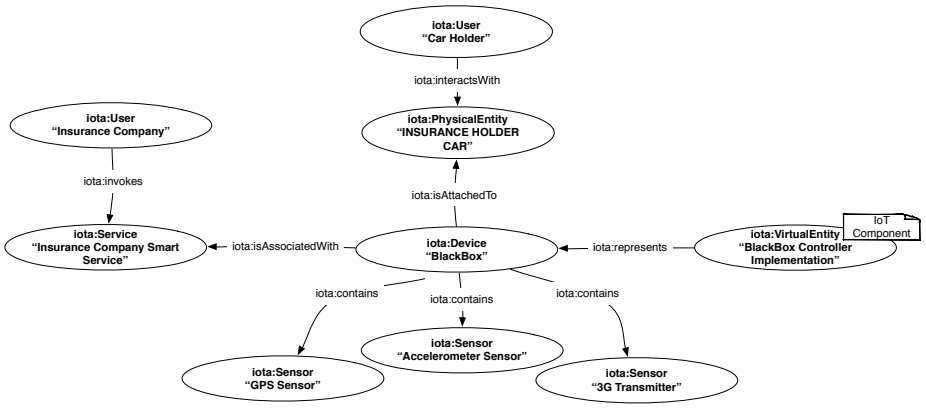

Fig. 3. Semantic model of car black box scenario modelled in IoTa Domain Model.

are being created generated (e.g. device generated information about the speed). We were able to align PROV-O with IoTa Domain Model so we can be more descriptive about IoT devices. Moreover by combination of these two semantic models we can better capture the creation,usage and flow of information in smart infrastructures.

Figure 4 shows an example of our framework being used to represent a in-car black box device. With the provenance support provided by PROV-O we are able to identify not only the high level concepts of IoT system, but specific data and relationships between them. In this case there is a process (*prov:Activity*) which used the GPS sensor to calculate current speed of vehicle and generated another *prov:Entity* containing calculated information about driver's speed. These information are then used by external process which calculates insurance premium. Note that both black box controller and insurance smart service acted on behalf of the same company. Participants during our design exercises highlighted the need to provide contact information about agents (individuals or organisations) responsible for certain devices and therefore we use the FOAF[8] ontology. The class foaf:Organization was defined as a subclass of prov:Agent.

3.1 T3 Model and Capability Inference Rules

As discussed earlier in this paper, PROV-O allows us to answer some of the issues related to transparency of IoT devices in smart environments, such as who is responsible for a device, what data is generated and where such data is transmitted. However, in PROV-O it is not possible to distinguish between personal and non-personal data, neither it is possible to describe the purpose behind the use of such data.

We designed an ontology (T^3 ontology) to support two important aspects of the reasoning required to infer and detect device capabilities. Firstly,the ontology

[8] http://xmlns.com/foaf/spec/

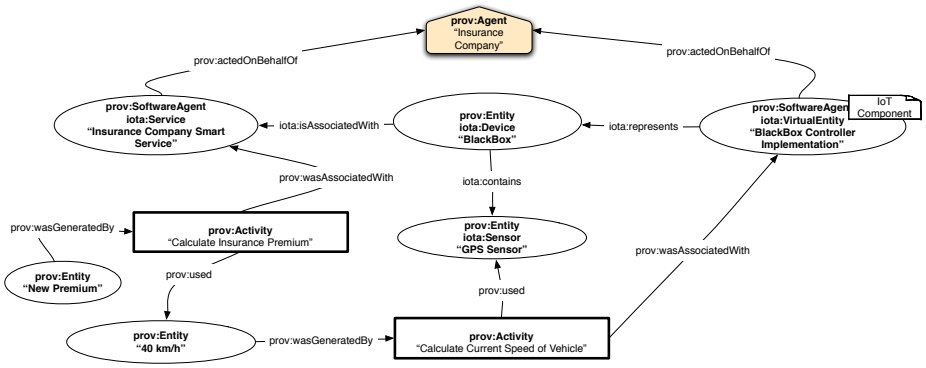

Fig. 4. Extract of the provenance from car black box scenario in combination with IoTa Domain Model

provides annotations over PROV-O concepts capturing the kind of information identified by our participants. These annotations include:

- The *ttt:PersonalData* class is used to identify if a *prov:Entity* represents data that can be personal to individual. Human-readable text of the data description can be attached to the *prov:Entity* via *ttt:description* property.
- The *ttt:purpose* property is used to provide an human readable explanation of why certain entities (described as personal data) are being generated or used. This property is associated with *prov:Activity* (process,task) via *prov:Usage* class.
- The *ttt:Capability* class defines different kinds of capabilities (e.g. *ttt:DataCollection*, *ttt:DataGeneration*, *ttt:DataSharing*,*ttt:DataGeneration*) that can be associated with *iota:Devices*. These associations (described by the *ttt:isCapableOf* property) are made on the basis of a number of inference rules described later in this section.
- Properties such as *ttt:securityDescription*, *ttt:deviceDesription*, *ttt:deviceName* were defined to describe and further support transparency in human readable format.

Secondly, in order to infer the capabilities of IoT devices using our ontological framework we can associate rules to specific classes of *ttt:Capability*. We make use of the SPIN ontology[9] to support the use of SPARQL to specify rules and logical constraints necessary to reason about capabilities. The rule is designed to traverse a PROV-O provenance graph starting from an instance of an *iota:Device*. To date, we implemented five inference rules, which are closely related to personal information: check if any IoT device or service is capable of generating automatic bills for a user, check if there are personal data being collected, rule to see if personal data are being used and why, detect if personal

[9] http://spinrdf.org/spin.html

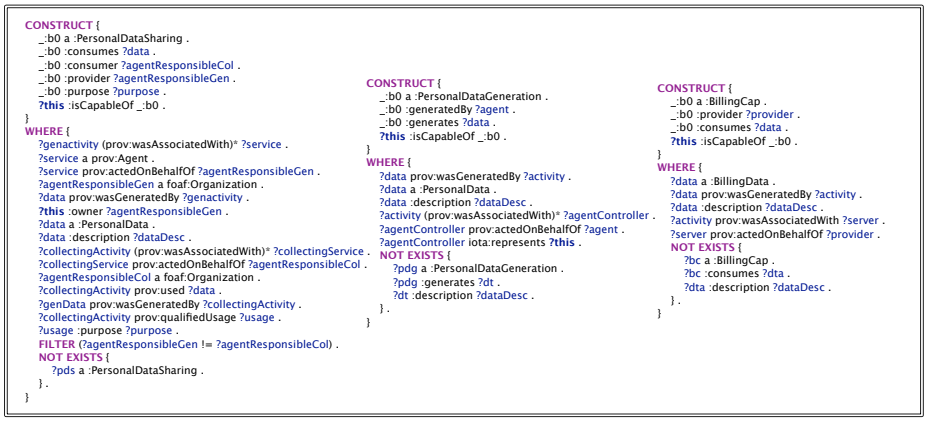

Fig. 5. Personal Data Sharing, Personal Data Usage and Billing Inference Rules

data are shared with third-party organisations, and detect what data is the device or service generating. Implementation of some of these rules can be seen in Figure 5.

If any of the rule activates, it constructs new inferred facts about the device in questions and forms them into capabilities (subclasses of *ttt:Capability*). If certain capability is already attached to the device (e.g. PersonalDataSharing) filtering service (see Section 2.3) makes sure it won't recreate the same one, unless there is new purpose or actor (e.g. there is a new agent the personal data are being shared with). In this case the same type of capability is attached to the device, but with changed properties.

If data are uniquely identified and the provenance of the devices and services associated with it is being captured and sent to our framework, system can track down how far certain information (e.g. personal data) flows in the whole IoT system or smart infrastructure. This will allow to identify all the agents, purpose and processes, that played any part in creation, generation, modification and aggregation of personally identifiable data.

We have also implemented rules that are used to determine what kind of provenance has been used to infer a specific device capability and if the capabilities are direct(on the device) or indirect(capability of external entities). However, this is out of the scope of this demo paper.

4 Demonstration Content

We organise this demo in two parts. The first part demonstrates how our mobile app can be used to query and visualise provenance information about IoT devices in the bus stop scenario. The second part of our demonstration explores a scenario, in which provenance is generated and sent to our system in real-time from IoT devices. We demonstrate how a user is notified about changes when the capabilities of an associated device change.

4.1 Bus Stops in Aberdeenshire

Near Field Communication (NFC) tags are being deployed on the bus stops in Aberdeenshire, Scotland UK so that user can retrieve real-time timetable information. These passive IoT devices embed a URL that is used to redirect and NFC-enabled smartphone to a service providing timetable information for a specific city location. Users may think that this service is provided by Aberdeenshire Council, but instead it delegates to third-party organisation (RSL Ltd). Moreover, both organisations are collecting certain kinds of personal data (e.g. IP Address, smartphone model and OS, and the version of the web browser users). Users deserve to know what data they are providing and to whom and they should also have an option to refuse to use the service. To this end, the Trusted Tiny Things tags and system have been deployed to more than 2300 bus stops in Aberdeenshire in 2014.

Demonstration content:

- For the purpose of this demonstration we built a replica bus stop. The NFC tag attached to the bus stop will be scanned using an NFC-enabled smartphone.
- We explain how the system checks if the device was previously accepted buy the owner of the smartphone. We demonstrate how the meta-data associated with the IoT device (bus-stop tag) and the meta-data associated with the timetable services is used in order to infer capabilities.
- Using the mobile app, we show what companies are associated with the device, who is the owner, manufacturer and list of capabilities, which will be explorable by clicking on them to find out further details about them. We will also demonstrate how the capability of the device can be accepted and recorded into the system.
- Finally, we will demonstrate how the system can be used for devices that have already been accepted where the app will show non intrusive dialog for few seconds informing the user about the time and date he trusted the device and how the user is immediately redirected to the service.

For a video presentation of this case study, please visit our website[10].

4.2 In-car Black box

The second part of the demonstration shows how our system can make use of real-time provenance data being generated by IoT devices. We simulate an insurance company that is installing tracker devices (black boxes) into their customer's vehicles so they can use the data collected to tailor the insurance premiums based on real driving behaviour. For this case study, we developed a replica black box and associated IoT services (insurance company service and a car manufacturer service) which can communicate and exchange information in a IoT ecosystem.

[10] http://t3.abdn.ac.uk

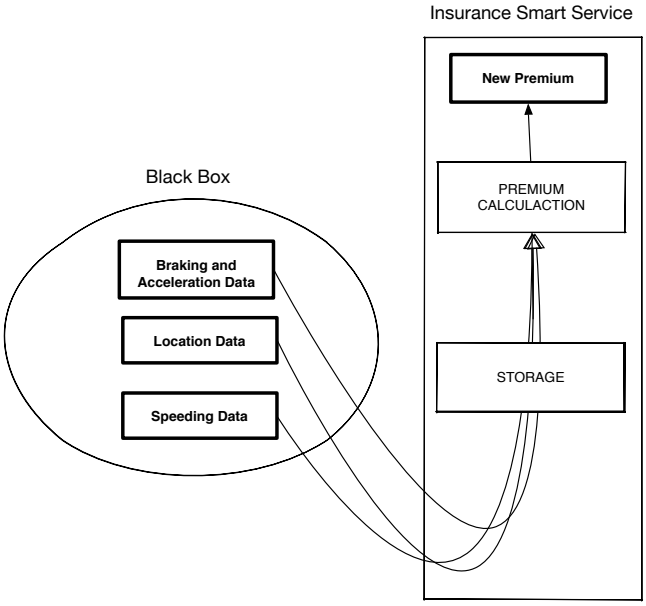

Fig. 6. Diagram illustrating the flow of data between the black box device and the insurance company service when calculating an insurance premiums

This scenario will be demonstrated in two stages:

In the first stage of this demonstration we will deploy a replica black box device to detect the location of a vehicle and the speed the vehicle is travelling. We will demonstrate how the insurance provider uses the information generated by the device to understand the driving patterns, current location and speed for premium calculation.

Using the Trusted Tiny Things mobile app we will scan the tag attached to the black box and visualise an initial set of capabilities. Such capabilities are defined based on the flow of data between the device and the supporting services. The flow of data can be seen in Figure 6.

In the second stage of this demonstration we will introduce new actor (car manufacturer service) to the IoT ecosystem. Previously, the insurance company was calculating premiums based on the information collected from the black box. In this scenario, the insurance company will begin to share accelerometer data with the car manufacturer's service in order to determine out how the policy holder's services their car.

We will illustrate how the new capabilities of the service are determined when the new service is activated. In this example, we will illustrate how the policy holder will be notified that personal data were shared with a third-party company. The flow of data associated with this scenario is illustrated in in Figure 7.

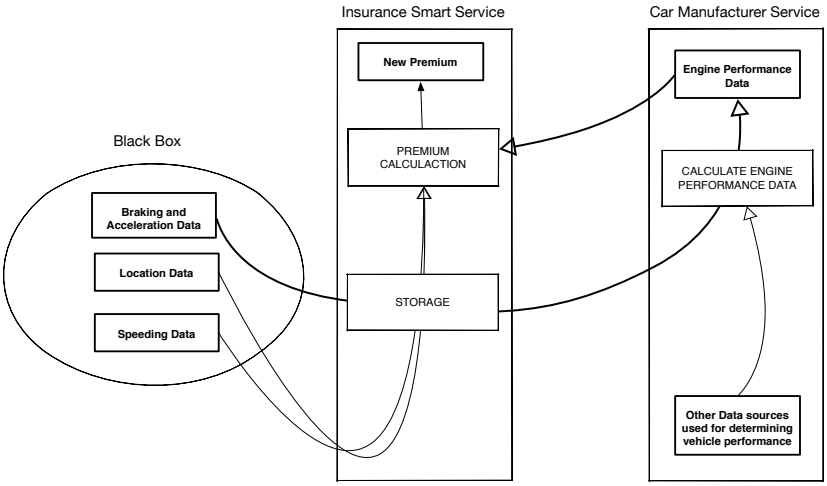

Fig. 7. Diagram illustrating the flow of data between the black box device and the insurance company service when a third party actor is involved

5 Acknowledgements

This research is supported by the Research Council UK Digital Economy IT as a Utility Network+ (EP/K003569/1) and the dot.rural Digital Economy Hub (EP/G066051/1).

References

1. Gusmeroli, S., Haller, S., Harrison, M., Kalaboukas, K., Tomasella, M., Vermesan, O., Vogt, H., Wouters, K.: Vision and challenges for realising the internet of things. In Friess, P., Guillemin, P., Sundmaeker, H., Woelfflé, S., eds.: Vision and Challenges for Realising the Internet of Things. European Commission (2010)
2. Hossain, E., Chow, G., Leung, V.C.M., McLeod, R.D., Mišić, J., Wong, V.W.S., Yang, O.: Vehicular telematics over heterogeneous wireless networks: A survey. Comput. Commun. **33**(7) (May 2010) 775–793
3. Pignotti, E., Beran, S., Edwards, P.: What does this device do? In: The First International Conference on IoT in Urban Space, October 27–28, 2014 Rome, Italy (to appear). (2014)
4. Bartoli, A., Hernández-Serrano, J., Soriano, M., Dohler, M., Kountouris, A., Barthel, D.: Security and privacy in your smart city. In: Proceedings of the Barcelona Smart Cities Congress. (2011)
5. Moreau, L.: The foundations for provenance on the web. Found. Trends Web Sci. **2**(2-3) (February 2010) 99–241
6. Bandara, A., Payne, T.R., de Roure, D., Clemo, G.: An ontological framework for semantic description of devices. McIlraith, S.A., Plexousakis, D., van Harmelen, F. (eds.) ISWC 2004. LNCS, vol. 3298, Springer, Heidelberg (2004)

7. Kotis, K., Katasonov, A.: An iot-ontology for the representation of interconnected, clustered and aligned smart entities. Technical report, VTT Technical Research Center, Finland VTT Technical Research Center, Finland (2012)
8. Bassi, A., Bauer, M., Fiedler, M., Kramp, T., van Kranenburg, R., Lange, S., Meissner, S.: Enabling Things to Talk. Springer (2013)
9. Serbanati, A., Medaglia, C.M., Ceipidor, U.B.: Building blocks of the internet of things: State of the art and beyond. Deploying RFID-Challenges, Solutions, and Open Issues, Dr. Cristina Turcu (Ed.), InTech (2011)
10. Haller, S.: The things in the internet of things. In: Internet of Things Conference, Tokyo, Japan, November 2010. (2010)